

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

<http://www1.ipdl.jpo.go.jp/PA1/result/detail/main/wAAAa23729DA410276449P1.htm> 01/09/04

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-276449

(43) 公開日 平成10年(1998)10月13日

(51) Int.Cl.⁸

H 0 4 N 9/79
5/91

識別記号

F I

H 0 4 N 9/79
5/91

K
N

審査請求 未請求 請求項の数 8 O L (全 9 頁)

(21) 出願番号 特願平9-79908

(22) 出願日 平成9年(1997)3月31日

特許法第30条第1項適用申請有り 平成8年11月22日
社団法人電子情報通信学会主催の「電子情報通信学会技
術研究報告会」において文書をもって発表

(71) 出願人 000004352

日本放送協会
東京都渋谷区神南2丁目2番1号

(72) 発明者 柴田 正啓

東京都世田谷区砧一丁目10番11号 日本放
送協会 放送技術研究所内

(72) 発明者 金 淵培

東京都世田谷区砧一丁目10番11号 日本放
送協会 放送技術研究所内

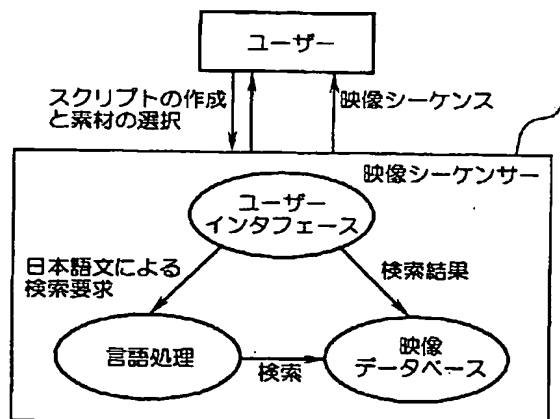
(74) 代理人 弁理士 谷 義一 (外1名)

(54) 【発明の名称】 映像編集システムおよび動画像分割方法

(57) 【要約】

【課題】 カメラワークの区切れおよび被写体のフレー
ムイン/アウトの時点境界とするセグメントに連続的
な映像を分割する。

【解決手段】 映像編集システムは映像のカラーヒスト
グラムの変化(非類似度曲線)を追跡し、変化の特性が
変わる時点(屈折点)を初期のセグメント境界点とす
る。映像編集システムは初期のセグメントの短い箇所や
平坦なものを隣接の初期のセグメントに統合して最終的
な分割セグメントを取得する。



1

【特許請求の範囲】

【請求項1】 動画像を複数のセグメントに分割する映像編集システムにおいて、
 動画像のフレームのカラーヒストグラムの変化の特性が変わる時点を検出する検出手段と、
 当該検出した時点を境界として、前記動画像を複数のセグメントに分割する分割手段とを具えたことを特徴とする映像編集システム。

【請求項2】 請求項1に記載の映像編集システムにおいて、前記検出手段は、前記動画像の全フレームについてフレームごとのカラーヒストグラムを求め、先頭フレームと各フレームの間のカラーヒストグラムの非類似度を求め、当該求められた非類似度を結ぶ非類似度曲線をスムージングして、屈折点を求め、該屈折点を前記カラーヒストグラムの変化の特性が変わる時点とすることを特徴とする映像編集システム。

【請求項3】 請求項2に記載の映像編集システムにおいて、前記分割手段は、前記屈折点を初期セグメントの境界とし、フレーム数が予め定められたしきい値以下の初期セグメントを直前の隣接の初期セグメントに統合することを特徴とする映像編集システム。

【請求項4】 請求項2に記載の映像編集システムにおいて、前記分割手段は前記非類似度曲線の単調変化部分の初期セグメント同士を統合することを特徴とする映像編集システム。

【請求項5】 動画像を複数のセグメントに分割する動画像分割方法において、
 動画像のフレームのカラーヒストグラムの変化の特性が変わる時点を検出し、
 当該検出した時点を境界として、前記動画像を複数のセグメントに分割することを特徴とする動画像分割方法。

【請求項6】 請求項5に記載の動画像分割方法において、前記動画像の全フレームについてフレームごとのカラーヒストグラムを求め、先頭フレームと各フレームの間のカラーヒストグラムの非類似度を求め、当該求められた非類似度を結ぶ非類似度曲線をスムージングして、屈折点を求め、該屈折点を前記カラーヒストグラムの変化の特性が変わる時点とすることを特徴とする動画像分割方法。

【請求項7】 請求項6に記載の動画像分割方法において、前記屈折点を初期セグメントの境界とし、フレーム数が予め定められたしきい値以下の初期セグメントを直前の隣接の初期セグメントに統合することを特徴とする動画像分割方法。

【請求項8】 請求項6に記載の動画像分割方法において、前記非類似度曲線の単調変化部分の初期セグメント同士を統合することを特徴とする動画像分割方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、映像を編集する映

2

像編集システムおよび動画像分割方法に関する。

【0002】

【従来の技術】 マルチメディア時代には、紙を媒体として流通してきたテキスト、図面、写真などの情報が、動画像や音声と統合され、デジタルデータとして流通するようになる。マルチメディアを処理するシステムは、映像（ここでは動画像と音声が結合した情報を指す）を信号として表示するだけでなく、表現された内容に基づいて映像を操作＝ハンドリングできることが重要である。表現された内容を記述し、コンピュータによる映像ハンドリングを可能にするデータをインデックスと呼ぶ。

【0003】 インデックスを付加された映像は、内容に基づいて検索することが可能になる。また、インデックスを使って構造化された映像（内容に基づいて分節化され、さらに分節間の関係が抽出された映像）は、時間軸に沿ったリニアな情報としての側面のほかに、全体を一覧したり、必要なセグメントのみを選択して見たり、あるいは再編集・再構成したりが可能な対話的な情報の側面を持つようになる。このことは、放送に対しても、映像制作を高度化・効率化したり、サービスをマルチメディア化するなどの大きな効果を及ぼすと考えられる。

【0004】 インデックスには、物理的記述のクラス（2次元的な構図、映像化された世界の3次元構造、カメラワークなど）と、言語的記述のクラス（被写体の名称、演技の内容、文脈など）の2つのクラスを考えることができる。これらのインデックスに基づいて映像を構造化し、インタラクティブなユーザーインタフェースを構築するための研究が数多く進められている（例えば、Glorianna Davenport, Thomas Aguiere Smith, Natali o Pincever : Cinematic Primitives for Multimedia. In IEEE Computer Graphics Applications, Vol.11, No.4, pp.67-74(1991)など）。このような研究に対し、我々は映像制作過程から取り出せる言語的記述に着目したアプローチを採っている。放送局などで映像が制作される過程では、制作者は映像内容の記述を取材メモあるいはシナリオという文書の形で残すことが多い。これらの内容記述は言語的記述のクラスに含まれ、被写体の名称、動作などの意味的な情報が自然言語によって表現されている。この制作者の内容記述は、制作された映像の記述としては最も重要なものと考えられる。我々は、この記述をインデックスとして扱うための手法と、このインデックスを使った自然言語インターフェースによる映像検索方式、およびインデックスに基づいて抽出した映像の時間的な階層構造を用いるブラウジング方式を提案している（下記参考文献（1）～（5）参照）。

【0005】（1）柴田正啓：“映像の内容記述モデルとその映像構造化への応用”、信学論(D-II), Vol. J78-D-II, No. 5, pp. 754-764(1995-5)

（2）Masahiro Shibata, Yeun-Bae Kim : C o n t e n

t-Based Structuring of Video Information. In Proc. of the IEEE International Conference on Multimedia Computing and Systems '96, Hiroshima, pp. 330-333 (1996-6)

(3) 柴田正啓、金淵培: “シーン記述に基づく映像の要約再生”、信学技報, PRU95-239, Vol. 95, No. 582, pp. 27-32 (1996-3)

(4) Yeun-Bae Kim, Masahiro Shibata: Content-Based Video Indexing and Retrieval - A Natural Language Approach -. In IEICE Trans. on Information and Systems, Vol. E79-D, No. 6, pp. 695-705 (1996-6)

(5) 金淵培、柴田正啓: “自然言語による映像シーケンサ”、TV学会誌, Vol. 50, No. 10, pp. 1585-1592 (1996-10)

上述のような映像編集システムでは、動画像を自動分割することが好ましい。動画像自動分割するための方法は、例えば、大辻、外村: 映像カット自動検出方式の検討、テレビ学会技報, Vol. 16, No. 43, pp. 7-12 (1992) において紹介されている。

【0006】

【発明が解決しようとする課題】 処理対象とする編集の素材となる動画像は、編集の結果として生じるつなぎ目を含まない。この連続して撮影された一続きの動画像を、その内容の変化に相当する被写体のフレームイン／フレームアウトの時点、およびカメラワークの始点／終点を検出して分割するためには、編集の結果として生じるつなぎ目を検出するための従来手法をそのまま適用することはできず、人間系すなわち、人間の目視確認に頼らざるを得なかった。

【0007】そこで、本発明の目的は、上述の点に鑑みて、カメラワークおよび被写体のフレームイン／アウトに基づく画面変化の区切れを自動検出して、動画像を分割することの可能な映像編集システム、およびその動画像の分割方法を提供することにある。

【0008】

【課題を解決するための手段】 このような目的を達成するために、請求項1の発明は、動画像を複数のセグメントに分割する映像編集システムにおいて、動画像のフレームのカラーヒストグラムの変化の特性が変わる時点を検出する検出手段と、当該検出した時点境界として、前記動画像を複数のセグメントに分割する分割手段とを具えたことを特徴とする。

【0009】請求項2の発明は、請求項1に記載の映像編集システムにおいて、前記検出手段は、前記動画像の全フレームについてフレームごとのカラーヒストグラムを求め、先頭フレームと各フレームの間のカラーヒスト

グラムの非類似度を求め、当該求められた非類似度を結ぶ非類似度曲線をスムージングして、屈折点を求め、該屈折点を前記カラーヒストグラムの変化の特性が変わる時点とすることを特徴とする。

【0010】請求項3の発明は、請求項2に記載の映像編集システムにおいて、前記分割手段は、前記屈折点を初期セグメントの境界とし、フレーム数が予め定められたしきい値以下の初期セグメントを直前の隣接の初期セグメントに統合することを特徴とする。

【0011】請求項4の発明は、請求項2に記載の映像編集システムにおいて、前記分割手段は前記非類似度曲線の単調変化部分の初期セグメント同士を統合することを特徴とする。

【0012】請求項5の発明は、動画像を複数のセグメントに分割する動画像分割方法において、動画像のフレームのカラーヒストグラムの変化の特性が変わる時点を検出し、当該検出した時点境界として、前記動画像を複数のセグメントに分割することを特徴とする。

【0013】請求項6の発明は、請求項5に記載の動画像分割方法において、前記動画像の全フレームについてフレームごとのカラーヒストグラムを求め、先頭フレームと各フレームの間のカラーヒストグラムの非類似度を求め、当該求められた非類似度を結ぶ非類似度曲線をスムージングして、屈折点を求め、該屈折点を前記カラーヒストグラムの変化の特性が変わる時点とすることを特徴とする。

【0014】請求項7の発明は、請求項6に記載の動画像分割方法において、前記屈折点を初期セグメントの境界とし、フレーム数が予め定められたしきい値以下の初期セグメントを直前の隣接の初期セグメントに統合することを特徴とする。

【0015】請求項8の発明は、請求項6に記載の動画像分割方法において、前記非類似度曲線の単調変化部分の初期セグメント同士を統合することを特徴とする。

【0016】本発明は、連続して撮影された一続きの動画像から、被写体のフレームイン／フレームアウトの時点、およびカメラワークの始点／終点を検出することを、画像フレームのカラーヒストグラムの時系列変化を測定することによって自動化する。カラーヒストグラムは、被写体の形状的な変化に比べ、被写体の画面への出入り（フレームイン／アウト）に敏感である。従って、画面内の被写体の入れ替わりや、カメラワークの有無を捉えることができる。

【0017】具体的には、先頭フレームを変化の基準として、先頭フレームのカラーヒストグラムと各フレームのカラーヒストグラムの非類似度を計算する。ここで、フレームのカラーヒストグラムは例えば次のようにして作成する。まず、カラーヒストグラムを構成する色空間を、例えばRGB空間をR軸、G軸方向に8分割、B軸方向に4分割した計 $8 \times 8 \times 4 = 256$ 個の各小空間

5

(ヒストグラムの欄)として定める。次に、フレーム画像を予め定められた大きさのブロック(例えば80×60ブロック)に分割し、各ブロックについて、そこに含まれる各画素の色情報に基づき、例えばその平均値をそのブロックの代表色とし、フレーム内全ブロックを対象にブロック単位でその代表色が256個の小空間の中のどの小空間に含まれるかを計算し、各小空間ごとの度数

$$d(t) = 1 - s(t)$$

$$s(t) = \frac{\sum_{k=1}^n \min(I_k^0, I_k^t)}{I}$$

ここで $s(t)$ はフレーム t と先頭フレームとの間の正規化されたカラーヒストグラム交差 ($0 \leq s(t) \leq 1$)

I_k^0 は先頭フレームのカラーヒストグラムの k 番目の欄の度数

I_k^t はフレーム t のカラーヒストグラムの k 番目の欄の度数

I はフレームに含まれるブロックの個数

n はヒストグラムの欄の総数 (上の例では $n = 256$)

【0019】非類似度 $d(t)$ は $0 \leq d(t) \leq 1$ の範囲の値をとり、両フレームが完全に同一の場合は d

(t) = 0 となり、両フレームに共通した代表色を含む小空間がない場合は $d(t) = 1$ となる。この非類似度の時間的変化が小さい部分は同一の被写体のカメラワークなしでの撮影に、変化があるが比較的緩やかな部分は 30 バンやズーム、移動などのカメラワークをしながらの撮影に、変化が大きい部分は主要な被写体のフレームイン／アウトに、それぞれ対応している。これらの区分に基づいて時間軸上で分割された部分動画像(セグメントと呼ぶ)の抽出のために、非類似度曲線の屈折点を求め、これらをセグメントの境界とする。

【0020】本発明ではセグメントの抽出は、非類似度曲線の屈折点を求め、これらをセグメントの境界とすることによって行う。屈折点を求める方法としては、例えば小山らの手法を使う(小山孝司、塩野充、真田英彦、40 手塚慶一: “細線化図形の屈折点検出の一手法”, 信学技報, PRL80-107, pp. 83-90(1980))。初期セグメントは見逃しを防ぐために、パラメータの設定によってやや過剰に検出した屈折点に基づいて求める。具体的には、スムージングに用いる移動平均の平均値算出点数を小さめに設定する。どの程度にするかは処理対象となる動画像の描写内容に応じて経験的に定める。

【0021】次に、過剰に検出した初期セグメントを、以下の手順で統合する。最初に、予め設定した閾値(例えば、5フレーム)以下の短小セグメントを直前の隣接

6

を計数してカラーヒストグラムを作成する。カラーヒストグラムの非類似度は、ヒストグラム間の交わり(交差)に基づいた尺度である。非類似度 $d(t)$ は次式により求める。

【0018】

【数1】

セグメントに統合する。次に、隣接するセグメントの非類似度曲線の変化率 = (セグメントの終点の非類似度 - セグメントの始点の非類似度) / (セグメントのフレーム数) の差が、2つの内の大きい方の変化率の例えば60%以下の隣接セグメント同士を単調変化部分として統合する。

【0022】例えば、あるセグメントの非類似度曲線の変化率が0.03で、次に続くセグメントの非類似度曲線の変化率が0.02である場合は、その差0.01は0.03の60%以下であるから両者を統合する。しかし、前者の変化率が0.03で後者の変化率が0.01である場合は、その差0.02は0.03の60%以下ではないので両者は統合しない。

【0023】最後に隣接するセグメントの非類似度曲線の1フレーム単位の変化率の平均値が、予め設定した閾値(例えば、0.03)以下の隣接セグメントどうしを平坦部分として統合する。そして、統合されなかったセグメントの境界を最終的な分割点として自動抽出する。

【0024】

【発明の実施の形態】以下、図面を参照して本発明の実施の形態を詳細に説明する。

【0025】図1に、自然言語による内容記述に基づく、素材映像のインデキシングの概要を示す。未加工の素材映像は、撮影された実世界の時間に対応する連続した時間軸を持つ映像シーケンスである。素材映像のインデキシングでは、「何が何をしている」といった内容に

7

基づいて、上記の時間軸上の該当部分を参照できるインデックスを作成する。制作者は、このインデックスを参照して必要な映像の部分を選択し、それらを番組としての時間軸の上に再配置する作業、つまり編集によって映像表現としての番組を完成させる。

【0026】インデックスの作成のためにはまず、内容に応じて素材映像を時間的に分節化する（セグメンテーション）。次に抽出した各分節（セグメント）について、その内容に対する記述情報を付加する。最後に、この記述情報をもとに、コンピュータ処理の可能なインデックスを生成する。通常の番組制作では、ビデオテープを試写しながら、制作者が人手によってセグメントを抽出し、その時間軸上の始点と終点をタイムコードやフレーム番号で、また、その内容を短い文で記述したメモを作成する。

【0027】本実施の形態の映像インデキシング方式では、このメモの電子化を前提に、入力された日本語の文をもとにインデックスを自動生成する。システムは、入力された内容記述の文を構文解析し、文を構成する文節間の係り受け関係を、主語を中心として表現するSubject-centered Dependency Structure(SDS)と呼ぶ木構造を生成する（上述の参考文献4参照）。このSDSをインデックスとすることにより、自然言語インターフェースによる効率のよい映像の内容検索が可能になる。本実施の形態で述べる映像編集システムは、この内容検索機能を応用している（上述の参考文献5参照）。一方映像の構造化のためには、SDSから抽出した内容記述文の基本要素を、時間軸上の存在範囲を持つ映像内容記述のための部品＝記述部品と考え、この記述部品の時間的布置＝スクリプトを生成する。スクリプト上の記述部品の分布をもとに映像の階層構造化を行うことにより、様々な形態の映像ハンドリングが可能になる（上述の参考文献1～3参照）。

【0028】以下では、インデックス生成の前提となる、電子化された映像内容記述メモの作成を支援するツールを提案する。このツールは、図1の中のセグメンテーションのステップを、映像フレームのカラーヒストグラムの時間的变化の情報を使って自動化する。カラーヒストグラムは、被写体の形状的な変化に比べ、被写体の画面への出入り（フレームイン／アウト）に敏感である。従って、画面内の被写体の入れ替わりや、カメラワークの有無を捉えることができる。映像の内容記述では「何が」という主語に相当する被写体（主被写体）が最も基本的な要素であり、主被写体のフレームイン／アウトのタイミングに基づくセグメンテーションは記述付与の有効な手がかりとなる。次に、抽出したセグメントに対する内容記述文の付与を行うが、映像情報、ことに放送番組の素材映像は意図に基づいて撮影されたものであり、内容記述にもこの意図が反映される必要がある。従って、提案するツールは撮影意図を把握しているユーザ

8

一、つまり制作者が使い、内容記述文はこの制作者がその撮影意図に基づいて入力するものとする。

【0029】図2に、映像内容記述メモ作成支援ツールのためのセグメンテーション法の手順を示す。セグメンテーション・プロセスの入力は連続して撮影された一続きの素材映像である。従って、番組として完成された映像とは異なり、カット点のような編集によって生じる不連続点を含まない。映像の主な変化要因は、カメラワークと被写体の動きである。素材映像はデジタル化され、適当な媒体に記録されているものとする。

【0030】手順ではまず、素材映像フレームのカラーヒストグラムの時系列変化を調べる。このためには、先頭フレームの変化の基準として、先頭フレームのカラーヒストグラムと各フレームのカラーヒストグラムの非類似度を計算する。この非類似度は、ヒストグラム間の交わり（交差）に基づいた尺度(Michael. Swain, DanaH. Ballard: Color Indexing. In Int. J. of Computer Vision, Vol. 7, No. 1, pp. 11-32(1991))であり、[0, 1]の範囲の値となる（検証実験では、画像サイズは800×600とし、ヒストグラムはRGB空間をR軸、G軸方向に8分割、B軸方向に4分割した計256個の欄について求めた）。カラーヒストグラムに基づくこの非類似度は、画面中で大面積を占める主被写体のフレームイン／アウトには急峻に反応して変化し、通常のカメラワークや被写体の変形に対しては緩慢に変化する特徴があることが、実験の結果確かめられた。

【0031】次に、求めた非類似度曲線をもとにセグメンテーションを行う。この過程では、非類似度の変化が一樣な区間をセグメントとして抽出する。変化が小さい部分は同一の被写体のカメラワークなしでの撮影に、変化があるが比較的緩やかな部分はパンやズーム、移動などのカメラワークをしながらの撮影に、変化が大きい部分は主被写体のフレームイン／アウトに、それぞれ対応していると考えられる。これらの区分に基づいたセグメントの抽出は、後段の内容記述文付与の重要な手がかりとなる。このようなセグメントの抽出のためには、非類似度曲線の屈折点を求め、これらをセグメントの境界とすればよい。屈折点を求めるアルゴリズムとしては小山らの手法があり（小山孝司、塩野充、真田英彦、手塚慶一：“細線化図形の屈折点検出の一手法”，信学技報，PRL80-107，pp. 83-90(1980)）、今回は非類似度曲線に移動平均によるスムージングを施した後、これを用いた。ステップ3で求める初期セグメントは“見逃し”を防ぐために、パラメータの設定によってやや過剰に検出した屈折点に基づいて求める。

【0032】次に、過剰に検出した初期セグメントを、図2のステップ4～6の手順で統合する。図3にこの過程の検証実験結果の例を示す。図中の小画像は各段階のセグメントを、その代表フレーム（時間軸上の中央のフレーム）で示している。この例では最終的に5つのセグ

メントが得られ、1、3、5番目のセグメントが人物を捉えたものであり、その間の2、4番目のセグメントはカメラワークが行われている部分となっている。

【0033】ここでは、初期セグメント検出と統合のためのパラメータの設定が問題となる。現在、種々の映像を使った検証実験を進めているが、非類似度曲線の振幅とスムージングのためのパラメータを調整するだけで、広い範囲の映像に対応できる見通しを得られている。図3は、カメラワーク（パン）による被写体のフレームイン／アウトの例であるが、カメラワークがない場合の被

写体自体の動きによるフレームイン／アウトも同様に検出できることが、実験の結果明らかになった。

【0034】図4に、前節で述べたセグメンテーションの機能を実現した映像内容記述エディタのGUI画面を模式的に表した図面を示す。エディタは、3つのウィンドウで構成される。図の右上のウィンドウは、素材映像のセグメンテーション結果（物理セグメントと呼ぶ）を表示する。ユーザーは、この中から画像アイコンと非類似度曲線を手がかりとして記述を付与するセグメント

（記述セグメントと呼ぶ）を指定して、エディタ・ウィンドウ（図の右下）に呼び出し日本語で映像内容を記述する。記述の登録された記述セグメントは、左側のウィンドウに列挙される。この作業の中でユーザーは、記述セグメントの境界を、エディタ・ウィンドウを使って適宜調整することができる。また、抽出された複数の物理セグメントをまとめて、1つの長い記述セグメントとすることが可能で、セグメントの包含関係に対応した階層的な記述も作成することができる。

【0035】以上の過程で、セグメント内容の文章表現はユーザーに委ねられているが、記述付与のガイドとなる枠組みが物理セグメントとして提示されることにより、メモ作成作業の大幅な効率化が期待できる。

【0036】次に、本実施の形態における素材映像の内容を普通の日本語の文で記述した構成表を用いる映像編集システム（映像シーケンサーと呼ぶ）について述べる（上述の参考文献5参照）。制作者は、所望の素材映像の内容を日本語文で記述し、個々の記述文を時間順に並べたリスト形式の構成表により映像シーケンスの構成を指定する。映像シーケンサーは、同じように日本語文を使ってインデキシングされた素材映像のデータベースを検索し、適合する部分映像（カット）を切り出し、順に再生することにより、映像シーケンスをシミュレーションする。この構成表は普通の日本語で素材映像とのリンクを取るため、構成表の可読性は極めて高くなる。この結果、映像シーケンサーを用いて、構成表上での映像シーケンスの構成の概念的な検討と、実際の映像を見ての具体的な検討を同時に効率的に行うことができる。

【0037】テレビ番組でもドラマの場合には、撮影前から綿密に構成を設計してシナリオ化し、シナリオに基づいて編集を行う。この他の大半の番組の編集では、実

際に撮影できた素材映像をもとに、映像としての構成を精密化していく過程が中心となる。ここで、素材映像の中のどの部分を、映像シーケンスの時間軸上のどこへ配置するかを粗く決める過程を粗編集と呼ぶ。粗編集では、構成の机上検討と素材映像の内容の確認を交互に繰り返し、つなぐカットとその時間順序を決定していく。その後の本編集では、つなぎの映像処理や音声処理を考慮して、タイミングを精密に決定する。映像シーケンサーは、上記の粗編集の過程を電子化する。

【0038】図5にシステムの構成を示す。システムは、ワークステーションの上に構築され、ユーザーインターフェース部、言語処理部、映像データベース部からなる。言語処理部は、NHK放送技術研究所で開発された日本語ニュース文用の構文解析システムを基に構成されている。また、映像データベース部には、日本語によってインデキシングされた素材映像がデジタル化されて蓄積されている。

【0039】ユーザーである制作者は、自分の想定するカットのシーケンスを普通の日本語で順次記述した表（構成表）をワークステーションの上に作成する（図6の左のウィンドウ）。構成表は通常のテキストエディタと同様な操作で、簡単に作成することができる。各行の文が所望のカットの内容を示しており、行の前後関係が映像シーケンスの構成に対応する。この段階では、構成表を参照してテキストベースで概念的に構成を練る。

【0040】この構成表を使った検討の過程では、適宜、実際に適当な素材映像のカットがあるか、複数の候補がある場合にはどのカットが最適かを、カットの前後関係も考慮して確認していく必要がある。また、素材映像の内容に応じて、構成を再検討することもある。本映像シーケンサーでは、言語処理部の機能により、構成表の各行の日本語文を映像データベースに対する検索質問として、適合する素材映像のカットを検索することができる。候補のカットが複数ある場合には、この中から最適のものを制作者が選び（図6の右のウィンドウが選択用のウィンドウ）、決定されたカットは構成表の中に画像アイコンとして表示される。制作者は、構成表の中の文の内容を自由に変更、追加、削除または順番を置き換えることによって新たな素材映像に簡単にアクセスすることができ、さらに文にカメラワークなどを条件として加えることで、カメラの操作とリンクされた、より適切な映像のシーケンスを得ることができる。

【0041】映像シーケンサーにおける素材映像の検索では、構成表の各文を素材映像のインデックスと比較対照することにより適合する映像カットを選び出す。このために、映像シーケンサーの言語処理部は制作者が作成した構成表の文をインデキシングと同様のプロセスでSDSに変換し、映像データベース部は素材映像のインデックスと構成表の文のSDSとの照合を行い、適合したSDSに対応する映像カット群を検索結果とする。この

11

手法には、以下の利点がある。

【0042】(1) 通常のキーワード検索で用いるBoolean Operator (AND/OR) だけでは表現できない「AがBを追いかける」といた助詞「が」、「を」、「へ」、「から」などの機能語による被写体間の関係を規定した検索に対応できる。

【0043】(2) 「Cが何かを食べている」のようなワイルドカード「何か」の使用を許すことによって柔軟な検索を可能にする。

【0044】(3) 特殊な検索言語が不必要であり、構成表の文を直接使用できる。

【0045】本実施の形態では、自然言語による映像ハンドリングに関連した技術として、素材映像のインデキシングと編集を支援するシステムについて述べた。放送などのように、ある意図のもとに撮影された素材映像を効果的かつ効率的に利用する必要がある場合には、言語的記述に基づくインデックスが有効である。このインデックスを映像の編集に利用する試みとしての映像シーケンサーについて、その基本的な仕組み、および自然言語による素材映像検索の方法について述べた。また、インデックスのもととなる映像内容記述の付与を支援するツールを提案した。このツールは、物理的な特性に基づいて素材映像を時間的に分節化することにより、ユーザーに記述付与のガイドを提供する。

【0046】今後、上記のツールで抽出したような物理的特性を、言語的記述に基づくインデックスと組み合わせ、さらに有用で、柔軟性のある映像ハンドリング環境を実現していく予定である。

【0047】以上、説明したように本実施の形態では、素材映像に対する言語的記述に着目し、これに基づく素材映像のインデキシングと、インデックスを使った編集について述べた。本実施の形態のインデキシングでは、言語的記述を素材映像に付与する作業を支援することができる。従来、この作業は、素材映像を収録した媒体(多くの場合はビデオテープ)を手動で操作し、試写を繰り返しながら映像の時間軸に沿ったセグメントを抽出し、これを自然言語で記述するといった手順で行われてきた。これに対し、本実施の形態では、この中のセグメント抽出の過程を映像処理の手法によって自動化する。

12

また、インデックスを使った編集システムでは、ユーザーは日本語の文で記述した構成表を元に、素材映像を検索し時間軸上への配置を行うことができる。

【0048】

【発明の効果】請求項1、5の発明ではカラーヒストグラムの非類似度の変化特性に基づきセグメントの境界を求めることにより従来では不可能であったカメラワークおよび人物等のフレームイン/アウトに基づく画面変化の区切りを検出することができる。このため、ユーザは映像の目視分割を要せず、大容量の映像の編集処理が容易となる。

【0049】請求項2、6の発明では、カメラワークの内容が異なる場合や人物等のフレームイン/アウトがある場合には、画面変化の区切り点が非類似度曲線上で屈折点として現れる点に着目し、これにより映像データから画面変化の区切り点を自動的に取得することができる。

【0050】請求項3、7の発明では、短いセグメントを隣接のセグメントに統合することにより、画面変化の途中をセグメント区切り点と誤検出した場合でも、セグメントの長さからその誤検出を見つけて、セグメント分割位置を修正することができる。

【0051】請求項4、8の発明では、非類似度曲線の平坦部分、すなわち、映像の内容が大きく変化しない部分を見つけセグメントを統合することにより異なる画面変化の区切り位置の検出精度を向上させる。

【図面の簡単な説明】

【図1】動画編集処理を示す説明図である。

【図2】セグメント分割の処理手順を示すフローチャートである。

【図3】セグメント分割の内容を示す説明図である。

【図4】映像編集用の表示画面を模式的に示す説明図である。

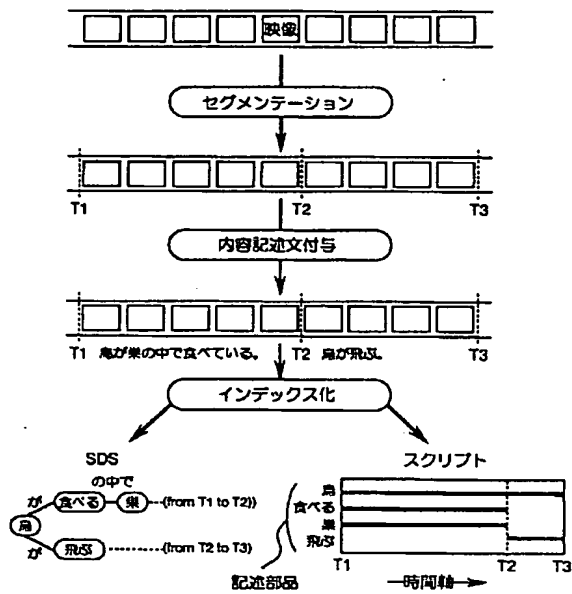
【図5】映像編集システムのシステム構成を示す構成図である。

【図6】映像編集システムの表示画面の一例を模式的に示す説明図である。

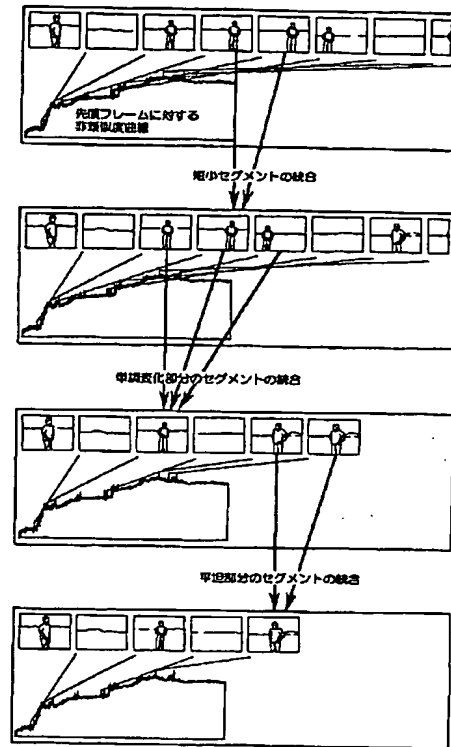
【符号の説明】

1 映像シーケンサー

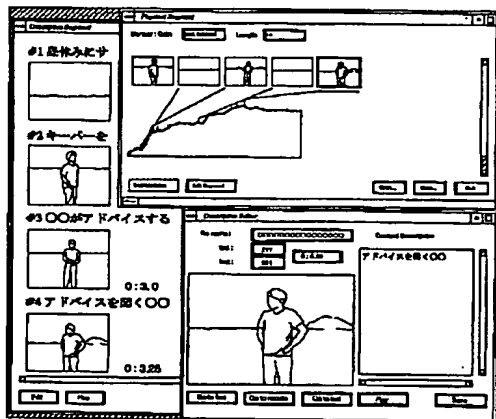
【図1】



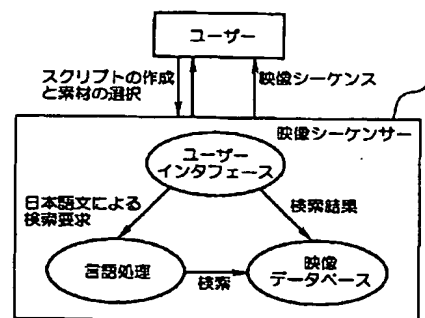
【図3】



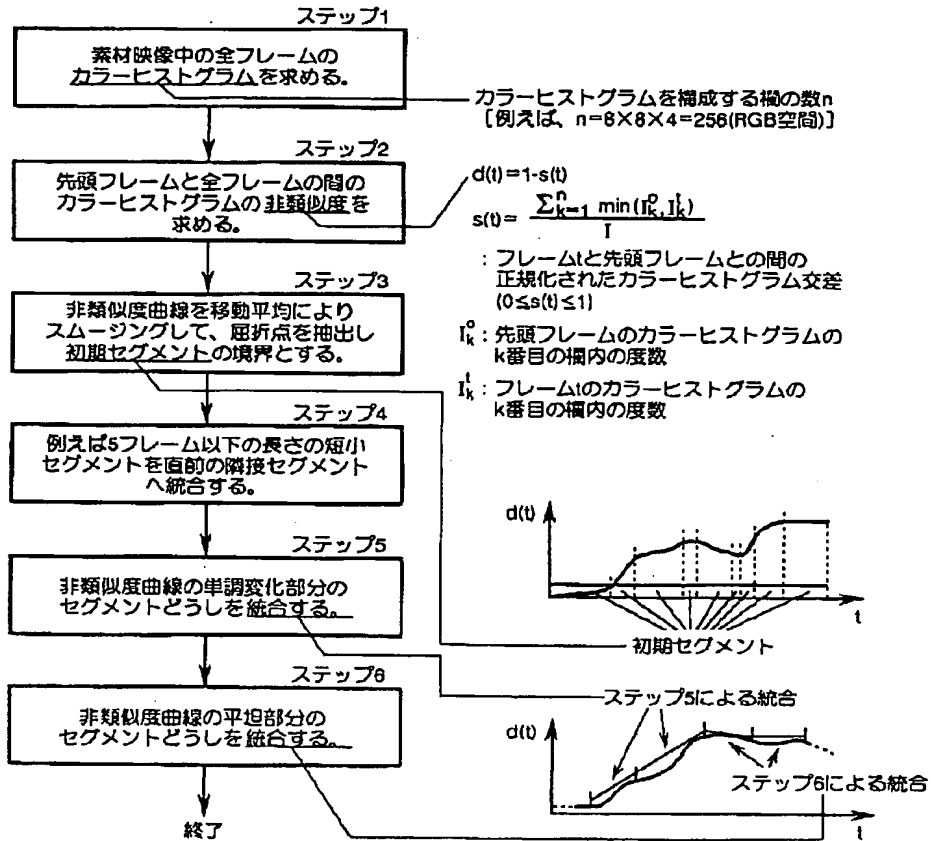
【図4】



【図5】



【図2】



【図6】

